LIS Overview and Related Training

Christa Peters-Lidard, Ph.D.[1] and the LIS Team

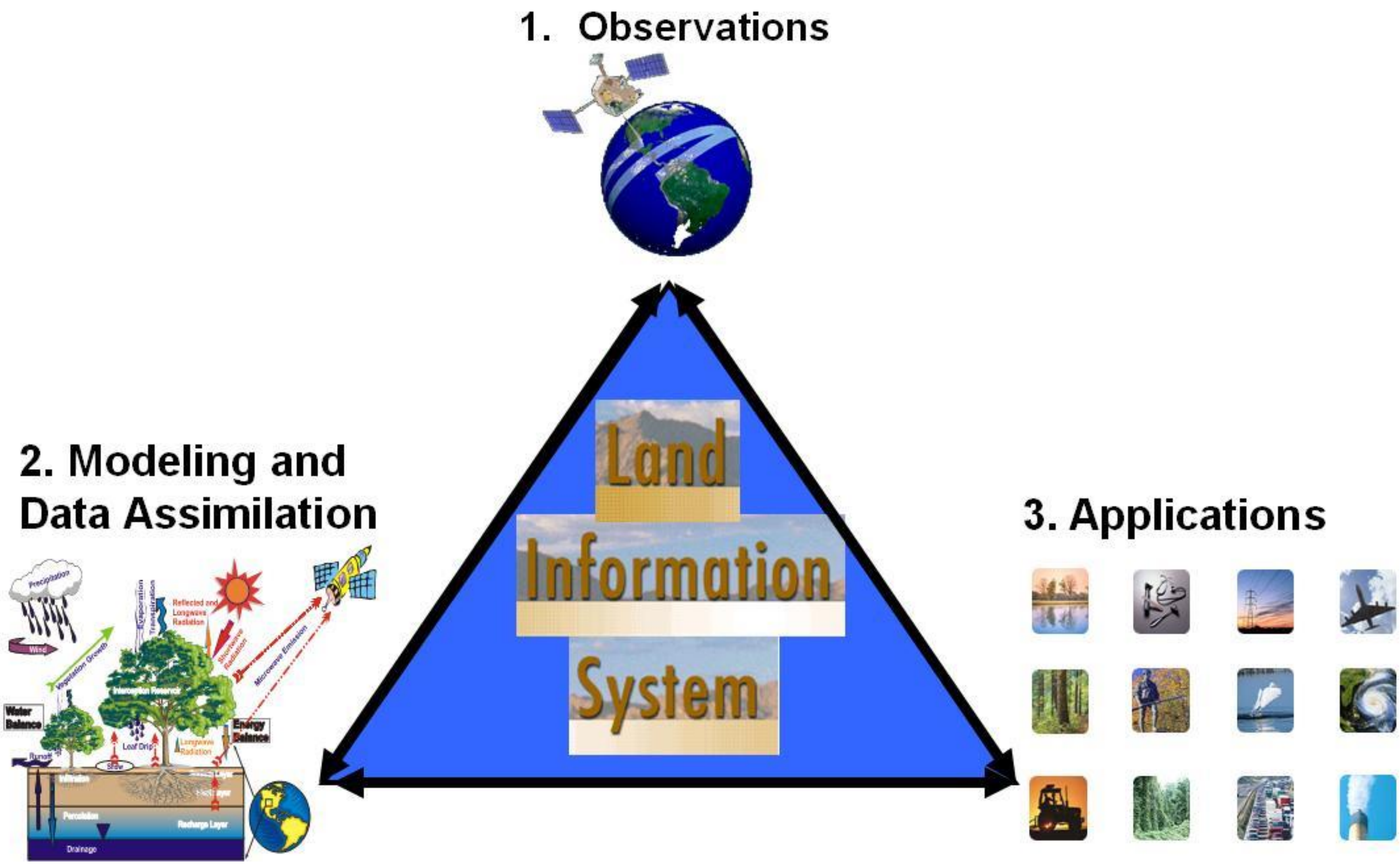
Sujay Kumar[2], David Mocko[2], Kristi Arsenault[2], Shugong Wang[2], Grey Nearing[2], Ken Harrison[3], Yuqiong Liu[3], Yudong Tian[3], Jim Geiger[1], Joe Santanello[1],

Partners:

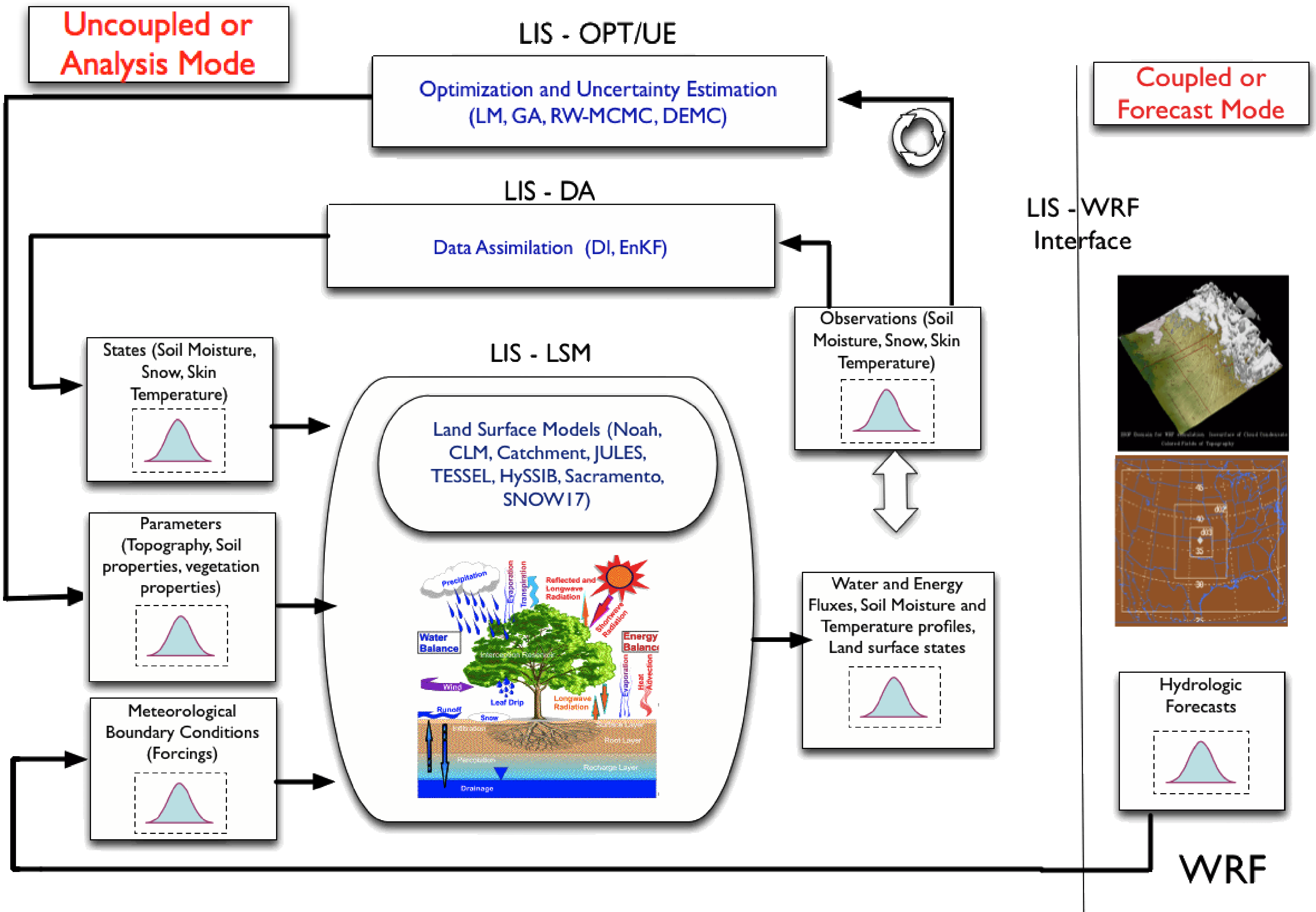
John Eylander[4], Jeff Cetola[5], Youlong Xia[6], Michael Ek[6], Chris Funk[7], Jim Verdin[7]

[1]Hydrological Sciences Laboratory, NASA/GSFC, Greenbelt, Maryland
[2]SAIC at Hydrological Sciences Laboratory, NASA/GSFC, Greenbelt, Maryland
[3]ESSIC at Hydrological Sciences Laboratory, NASA/GSFC, Greenbelt, Maryland
[4]Army Cold Regions Research & Engineering Lab, Engineer Research & Development Center
[5]16th Weather Squadron (16WS/WXE), 2nd Weather Group (Air Force Weather Agency), Offutt AFB, NE
[6]Environmental Modeling Center, NCEP/NOAA, Camp Springs, Maryland
[7]USGS, Eros Data Center, Sioux Falls, South Dakota

LIS Integrates Observations, Models and Applications to Maximize Impact

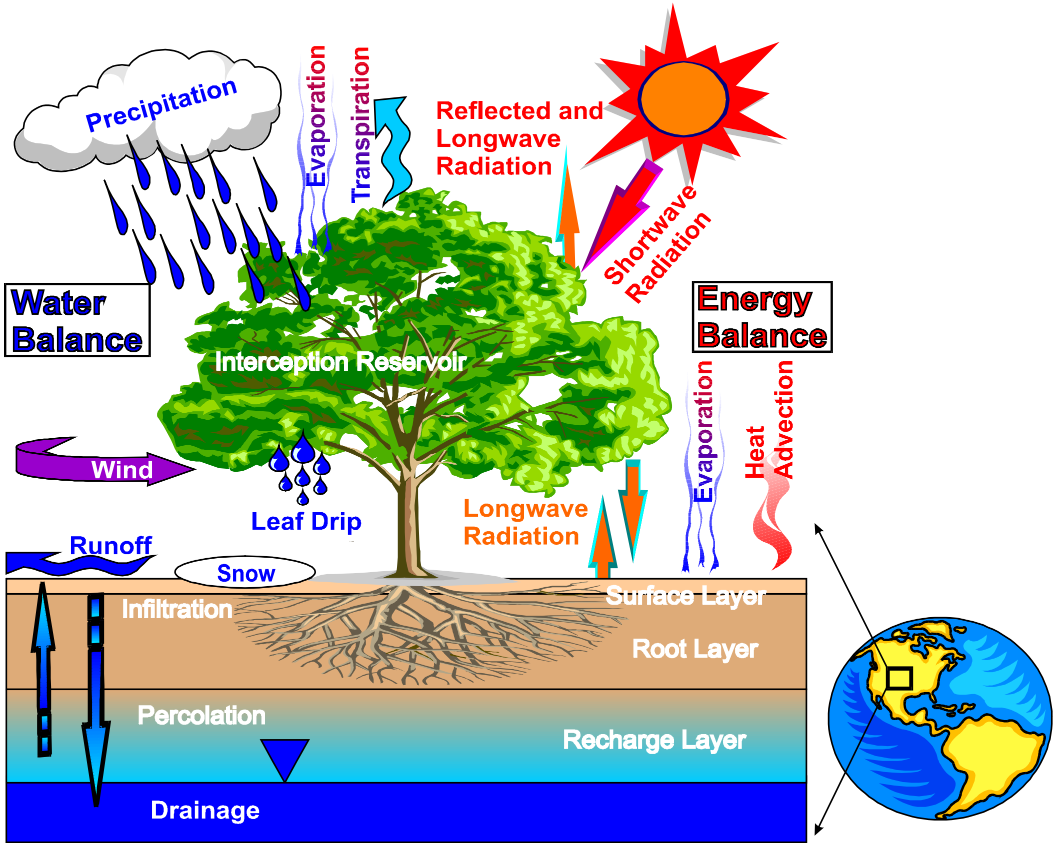


The Land Information System (http://lis.gsfc.nasa.gov)

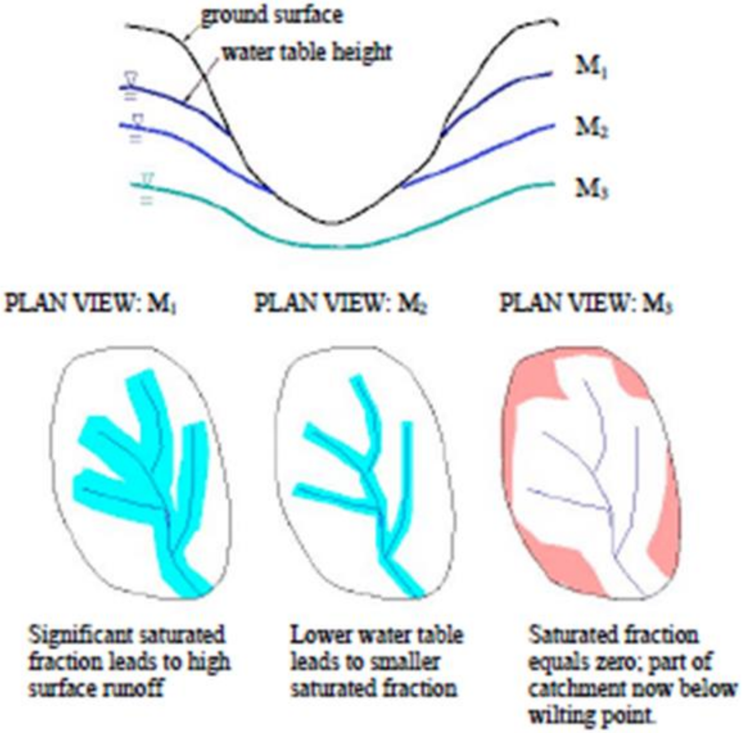


LSM Hydrology: Water and Energy Balance

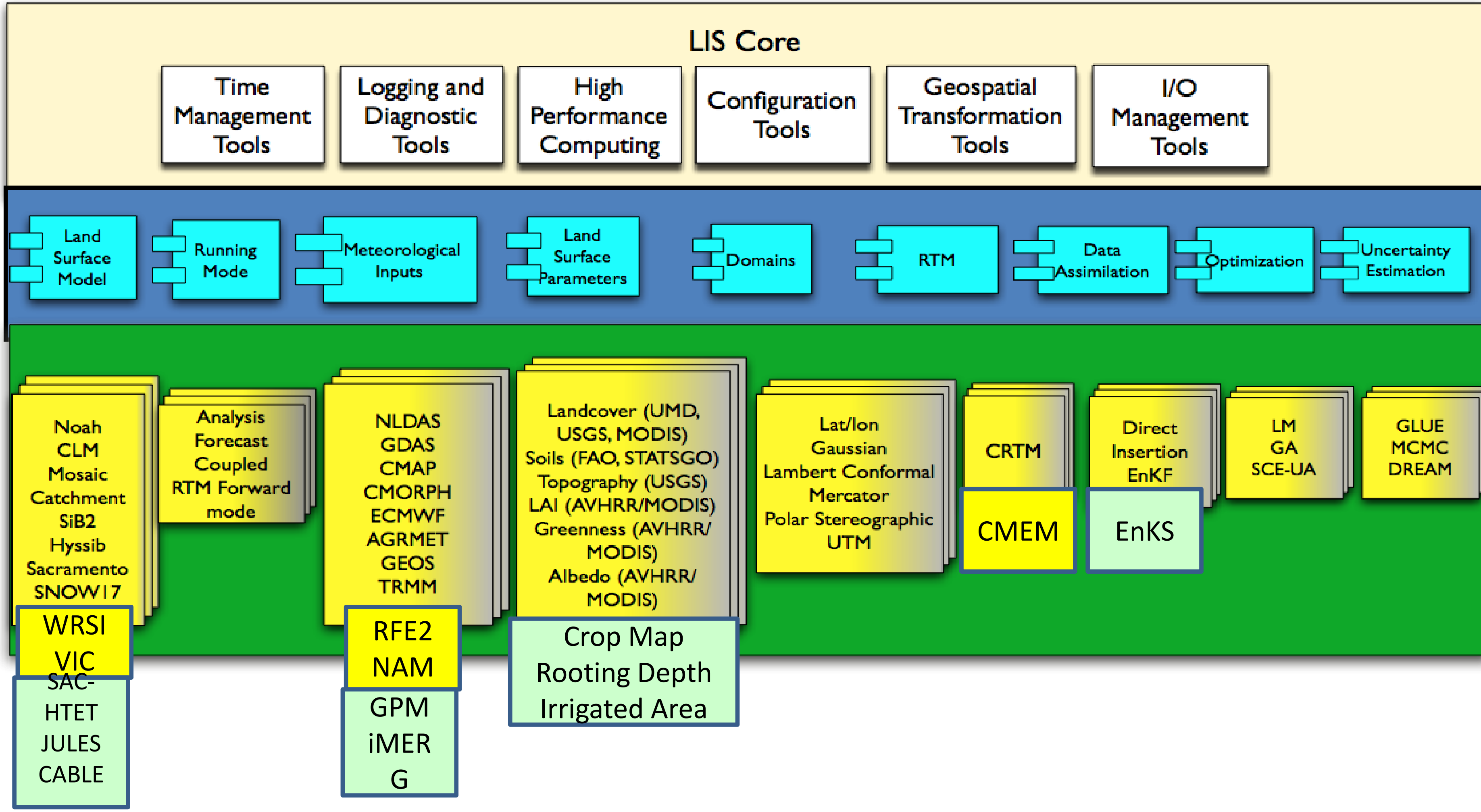
e.g., Noah, CLM, VIC, TESSEL, JULES, HySSIB



Catchment



Land Information System Architecture



Kumar, S. V., C. D. Peters-Lidard, Y. Tian, P. R. Houser, J. Geiger, S. Olden, L. Lighty, J. L. Eastman, B. Doty, P. Dirmeyer, J. Adams, K. Mitchell, E. F. Wood and J. Sheffield, 2006. Land Information System - An Interoperable Framework for High Resolution Land Surface Modeling. *Environmental Modelling & Software,* Vol. 21, 1402-1415.

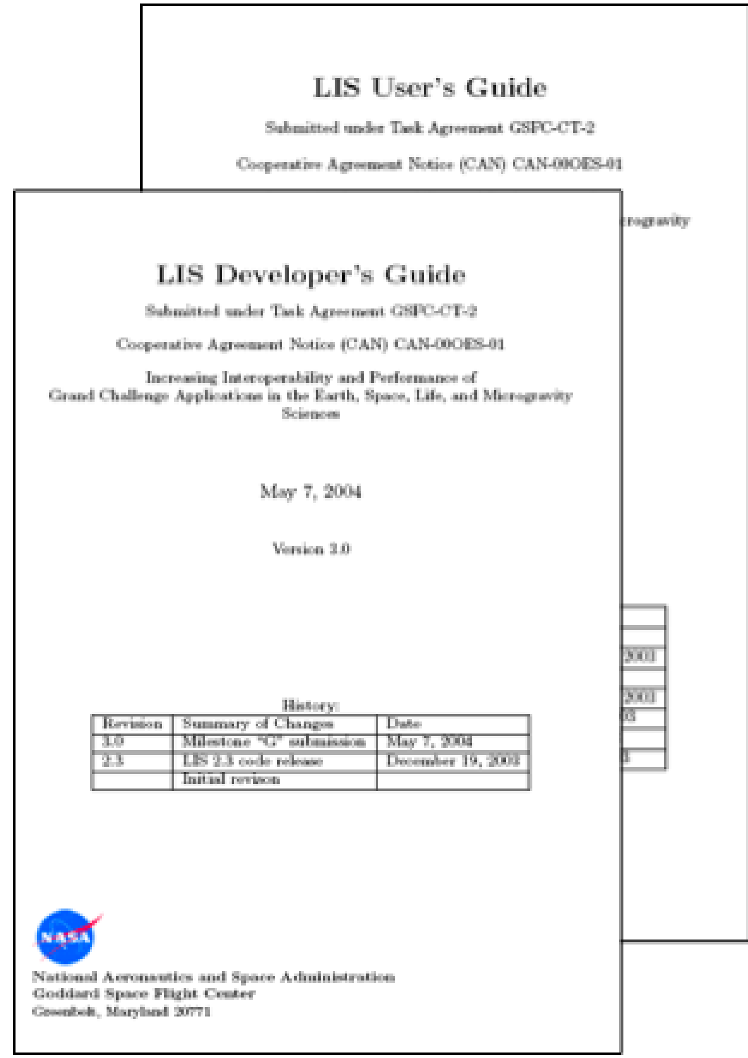
Peters-Lidard , C.D., P.R. Houser, Y. Tian, S.V. Kumar, J. Geiger, S. Olden, L. Lighty, B. Doty, P. Dirmeyer, J. Adams, K. Mitchell, E.F. Wood and J. Sheffield, 2007: High-performance Earth system modeling with NASA/GSFC's Land Information System. *Innovations in Systems and Software Engineering.* 3(3), 157-165. [DOI:10.1007/s11334-007-0028-x](DOI:10.1007/s11334-007-0028-x)

LIS on-line tutorial and support

http:://lis.gsfc.nasa.gov http:://modelingguru.nasa.gov

LIS Documentation

- User’s guide
  - –Step-by-step instructions on how to build the LIS code
- Developer’s guide
  - –Instructions on how to bring in new functionalities (LSMs, forcing schemes, Data Assimilation, parameter data, etc.)
- Reference manual



LIS6.2 Software Updates

- Includes VIC 4.1.1.
- Includes CABLE 1.4b --- restricted distribution.
- Includes Catchment F2.5.
- Includes Noah 3.3.
- Includes SiB2.
- Includes WRSI.
- Support for North American Mesoscale Forecast System (NAM) "242 AWIPS Grid -- Over Alaska" product.
- Support for USGS potential evapotranspiration (PET) data (for use in WRSI).
- Support for Climate Prediction Center's (CPC) Rainfall Estimates version 2 (RFE2) daily precipition (for use in WRSI).
- Support to apply lapse-rate correction to bottom temperature field (for use in Noah).

LIS6.2 Software Requirements

- ☑ Fortran 90/95 compiler (g95 will not work for LIS5.0)
  - ☐ preferred : intel, pgi, lahey, absoft
- ☑ C compiler
- ☐ MPI - if parallel processing capability is desired
- ☑ Earth System Modeling Framework (ESMF)
  - ☐ 3.1.0r - for LIS 6.x
- ☐ LIS supports Grib1, NETCDF, HDF formats
  - ☐ Grib1 - mandatory, NETCDF, HDF optional

LIS Land Data Assimilation Objectives

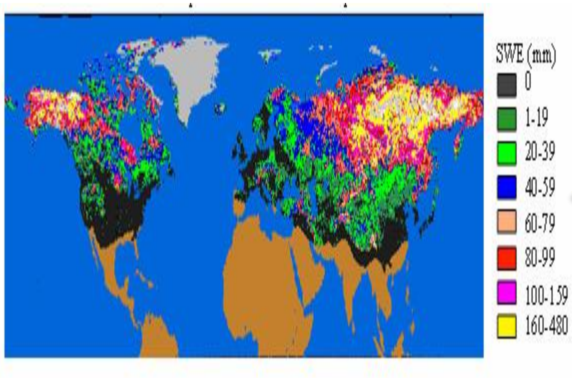
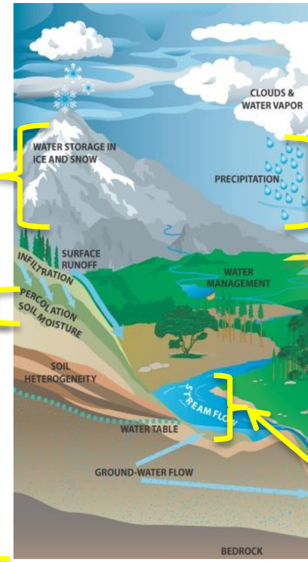


Figure 1: Snow water equivalent (SWE) based on Terra/MODIS and Aqua/AMSR-E. Future observations will be provided by JPSS/VIIRS and DWSS/MIS.



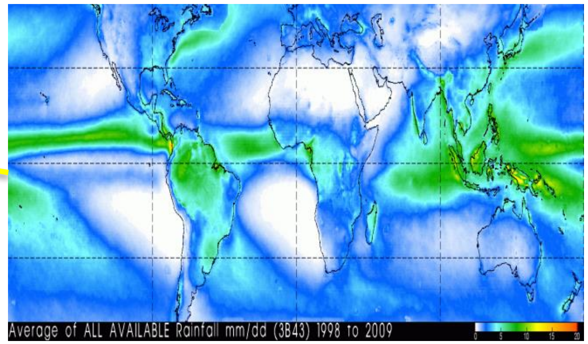


Figure 2: Annual average precipitation from 1998 to 2009 based on TRMM satellite observations. Future observations will be provided by GPM.

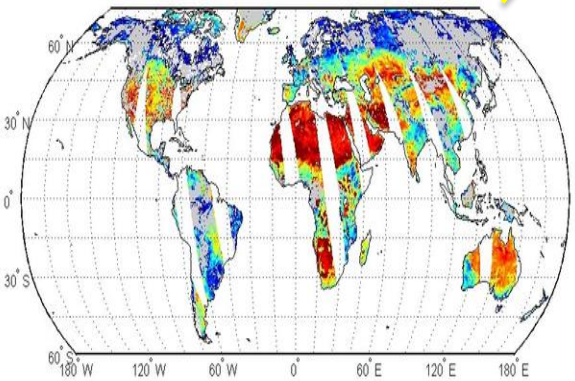


Figure 3: Daily soil moisture based on Aqua/AMSR-E. Future observations will be provided by SMAP.

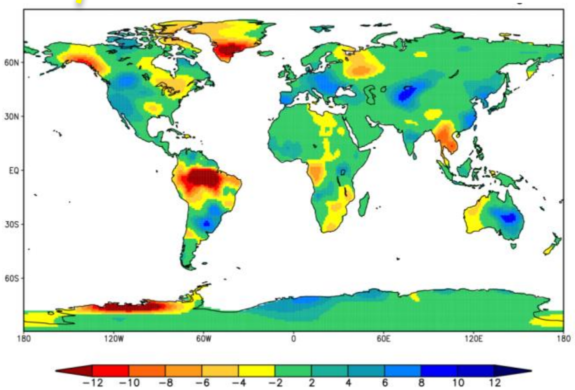


Figure 4: Changes in annual-average terrestrial water storage (the sum of groundwater, soil water, surface water, snow, and ice, as an equivalent height of water in cm) between 2009 and 2010, based on GRACE satellite observations. Future observations will be provided by GRACE-II.

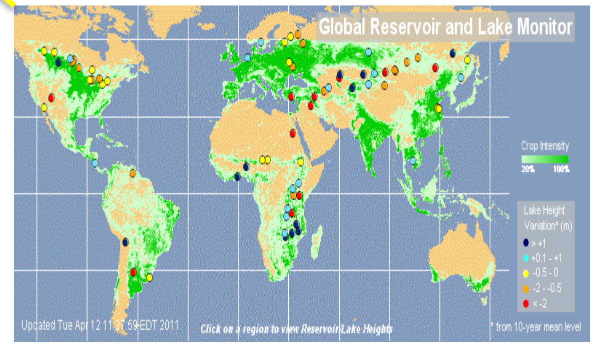
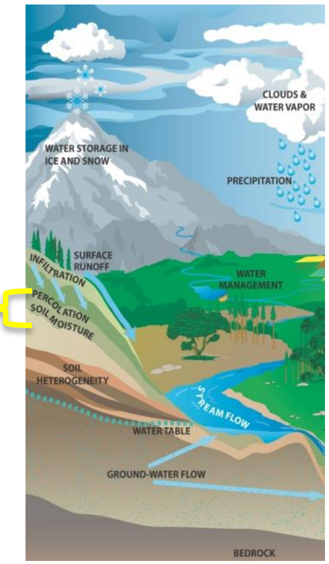


Figure 5: Current lakes and reservoirs monitored by OSTM/Jason-2. Shown are current height variations relative to 10-year average levels. Future observations will be provided by SWOT.

Soil Moisture Data Assimilation

Experimental Setup:
- Domain: CONUS, NLDAS
- Resolution: 0.125 deg.
- Period: 1979-01 to 2012-01
- Forcing: NLDASII
- LSM: Noah 3.3



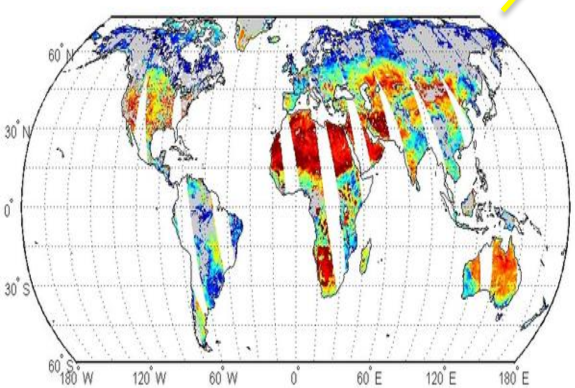
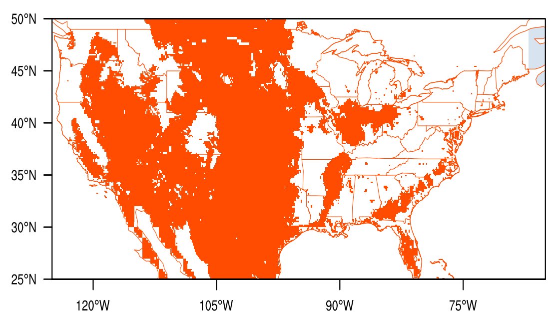


Figure 3: Daily soil moisture based on Aqua/AMSR-E. Future observations will be provided by SMAP.

Data Assimilation:
- AMSR-E LPRM (Owe et al., 2008; Peters-Lidard et al., 2011) 2002-2011
- ESA ECV (Liu et al., 2012; Wagner et al., 2012) 1978-2011
- Flags: light and moderate vegetation, precipitation, snow cover, frozen ground, RFI
- The observations are scaled to the LSM's climatology using CDF matching
- 12-member ensemble
- A spatially distributed observation error standard deviation (between 0.02-0.12 m3/m3)



Effective mask of the locations where at least one year of data is assimilated

Evaluation of LIS+DA outputs

Soil moisture:

USDA Soil Climate Analysis Network (SCAN); 37 stations chosen after careful quality control (used for evaluations between 2000-2011)

Four USDA ARS experimental watersheds (“CalVal” sites) (used for evaluations between 2001-2011)

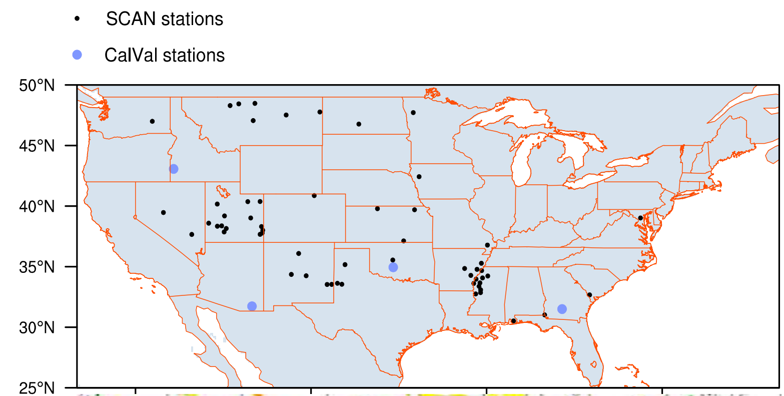
Streamflow:

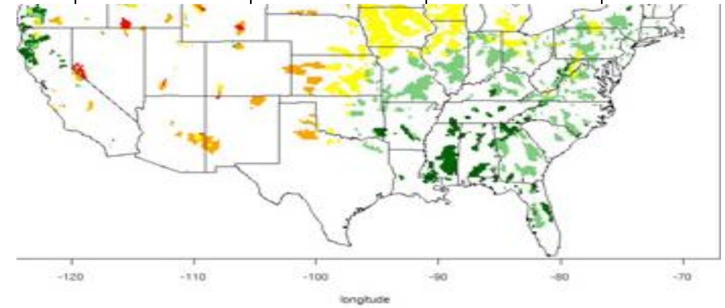
Gauge measurements from 961 unregulated USGS streamflow stations (1981-2011) from Xia et al., 2012.

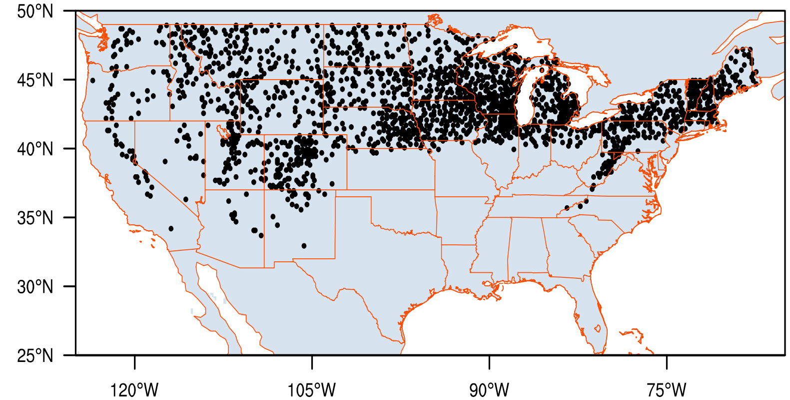
Snow depth:

Global Historical Climate Network (GHCN) – used for evaluations between 1979-2011.

Canadian Meteorological Center (CMC) daily snow depth analysis – used for evaluations between 1998-2011.

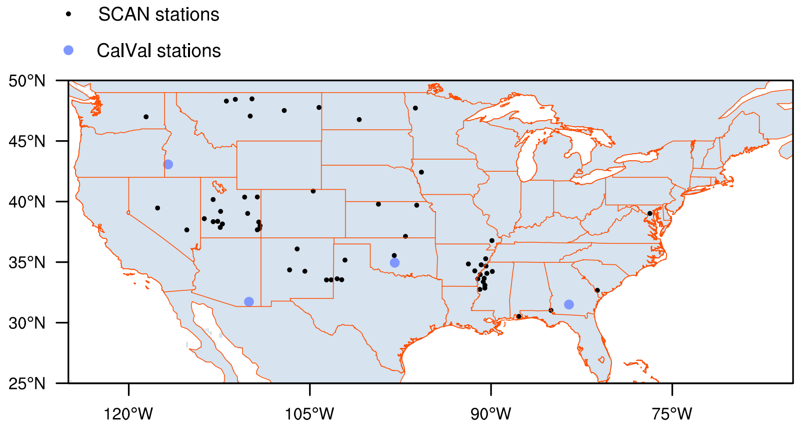






All model verifications and analysis generated using the Land surface Verification Toolkit (LVT; Kumar et al. 2012)

Soil moisture DA : Evaluation of soil moisture fields



ARS CalVal (surface soil moisture)	Open loop (no DA)	LPRM DA
Anomaly R	0.84 +/- 0.02	**0.86 +/- 0.02**
Anomaly RMSE (m3/m3)	0.021 +/- 0.001	**0.019 +/- 0.001**
ubRMSE (m3/m3)	0.024 +/- 0.002	**0.022 +/- 0.002**

Statistically significant improvements in surface soil moisture and root zone soil moisture as a result of soil moisture DA

Anomaly R increases, Anomaly RMSE reduces and unbiased RMSE reduces with LPRM assimilation.

SCAN (surface soil moisture)	Open loop (no DA)	LPRM DA
Anomaly R	0.67 +/- 0.02	**0.67 +/- 0.02**
Anomaly RMSE (m3/m3)	0.037+/- 0.002	**0.036 +/- 0.002**
ubRMSE (m3/m3)	0.043 +/- 0.003	**0.041 +/- 0.003**

SCAN (root zone soil moisture)	Open loop (no DA)	LPRM DA
Anomaly R	0.60 +/- 0.02	**0.59 +/- 0.02**
Anomaly RMSE (m3/m3)	0.032 +/- 0.002	**0.030 +/- 0.002**
ubRMSE (m3/m3)	0.041 +/- 0.003	**0.039 +/- 0.003**

Soil Moisture DA: Evaluation of improvements in streamflow simulation

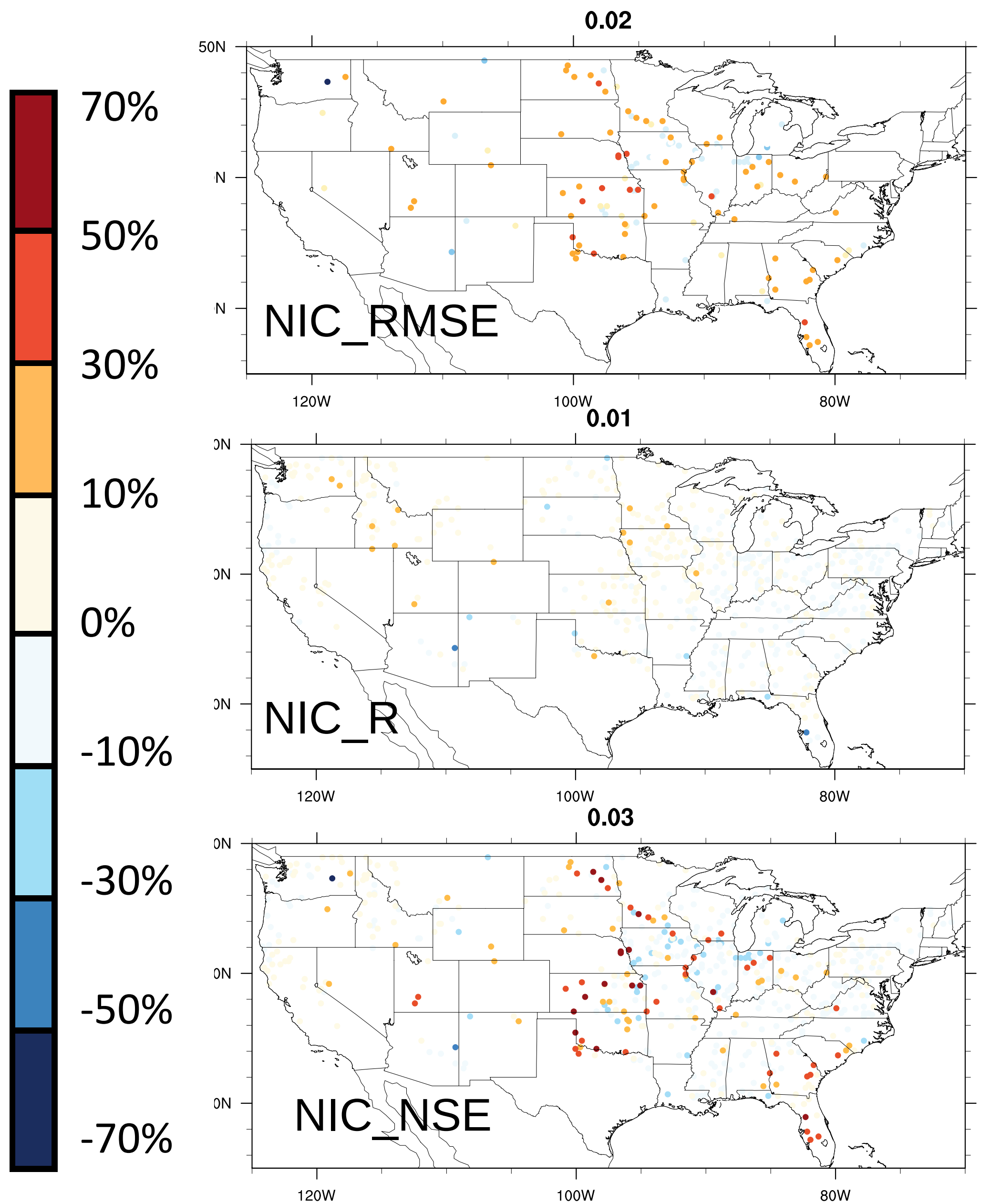
The improvements are expressed using a Normalized Information Contribution (NIC) metric that measures the skill improvement from DA as a fraction of the maximum possible skill improvement

$$NIC_{RMSE} = \frac{(RMSE_o - RMSE_a)}{RMSE_o}$$

$$NIC_R = \frac{(R_a - R_o)}{(1 - R_o)}$$

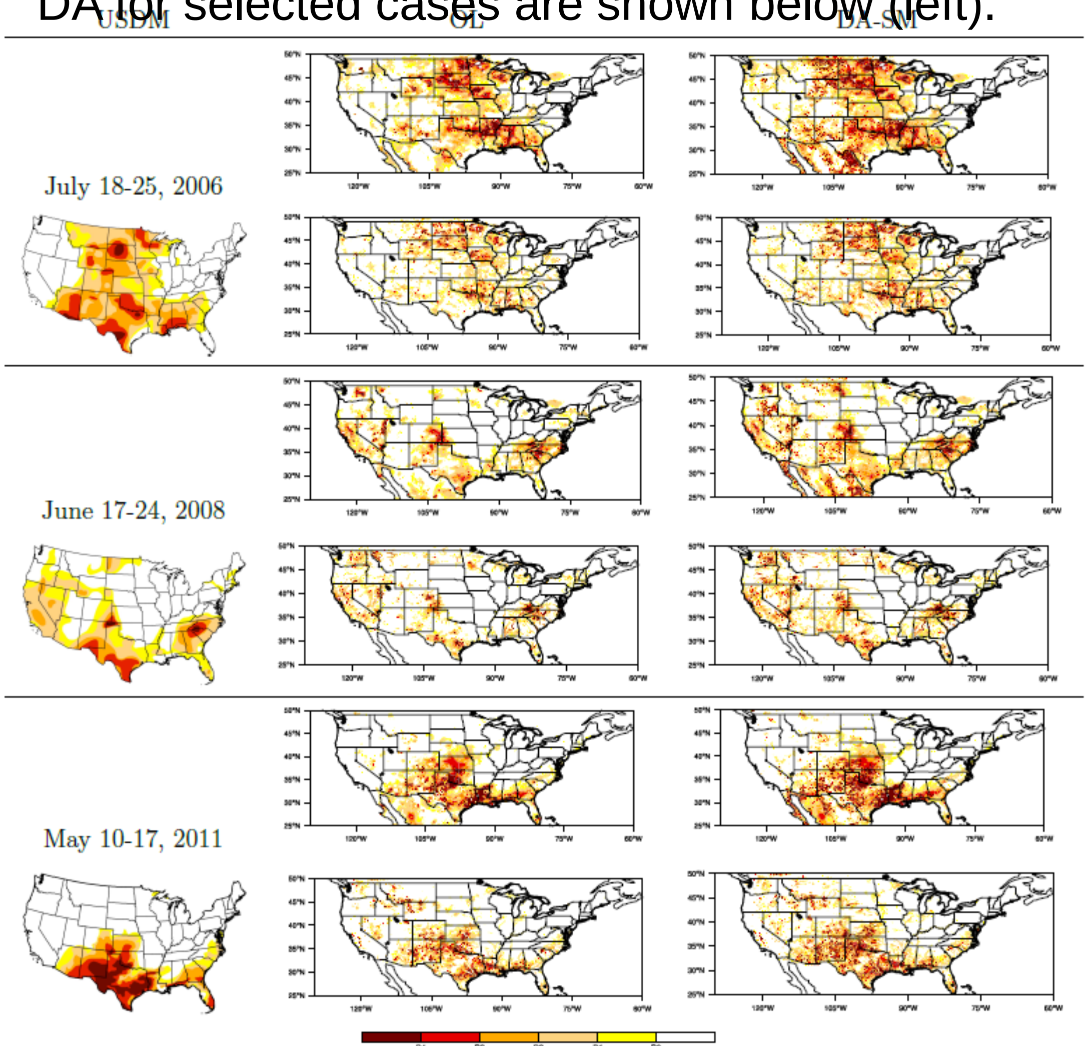
$$NIC_{NSE} = \frac{(NSE_a - NSE_o)}{(1 - NSE_o)}$$

Overall improvements in all skill metrics (RMSE, R and NSE) are observed in streamflow estimates after data assimilation



Soil moisture DA: Evaluation against USDM drought percent

Comparisons are made against percentage of area under different drought conditions from the USDM (2000-2011) for each Region. An example of the RMSE (%) of the D0 drought conditions is shown below (right; **bold** is statistically significant). Changes in the root zone soil moisture percentiles from soil moisture DA for selected cases are shown below (left).



RMSE (%) of simulated drought areal extent as compared to USDM analysis

Region (D0 area)	Open Loop (no DA)	Soil moisture DA
South	15.7	**12.8**
Southeast	22.6	**20.9**
Northeast	17.0	17.0
Midwest	10.6	***11.5***
High Plains	32.1	**27.6**
West	30.3	**25.9**

Snow Data Assimilation

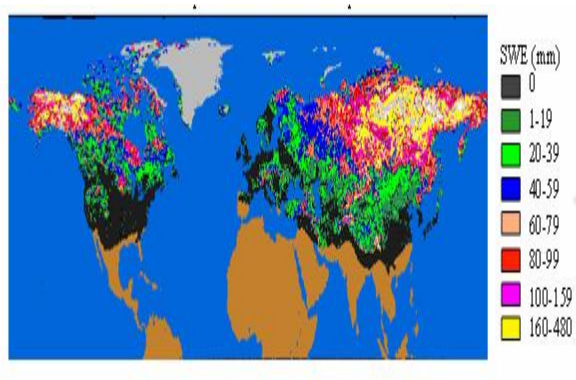
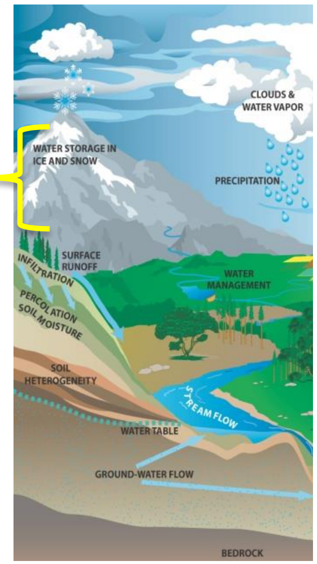


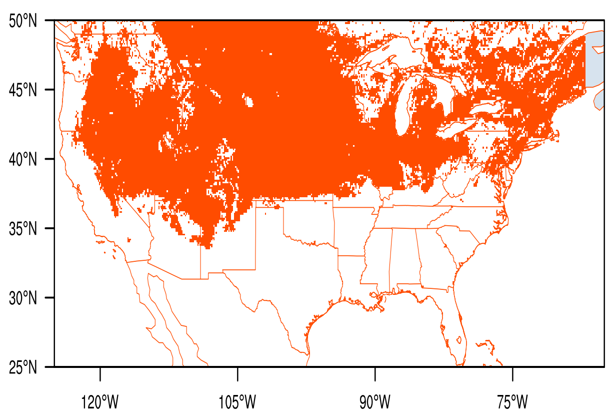
Figure 1: Snow water equivalent (SWE) based on Terra/MODIS and Aqua/AMSR-E. Future observations will be provided by JPSS/VIIRS and DWSS/MIS.

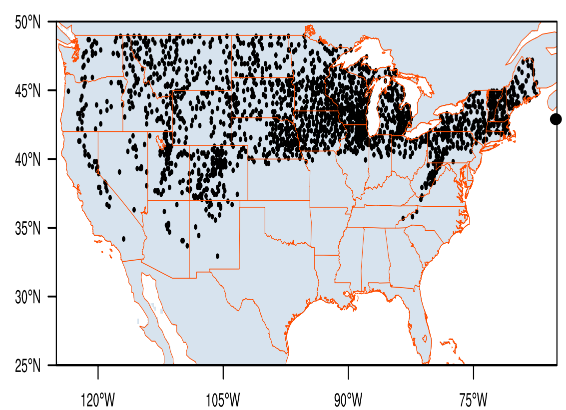


Data Assimilation:

- SMMR (spans 1978-1987), SSM/I (spans 1987-2002) and AMSR-E (spans 2002-2011); SMMR and SSM/I retrievals are based on the Chang et al. (1987) and AMSR-E retrievals are based on the improved retrieval algorithm from Kelly et al. (2009).
- The snow depth retrievals are corrected using the in-situ measurements from the Global Historical Climate Network (GHCN).

Effective mask of the locations where at least one year of data is assimilated

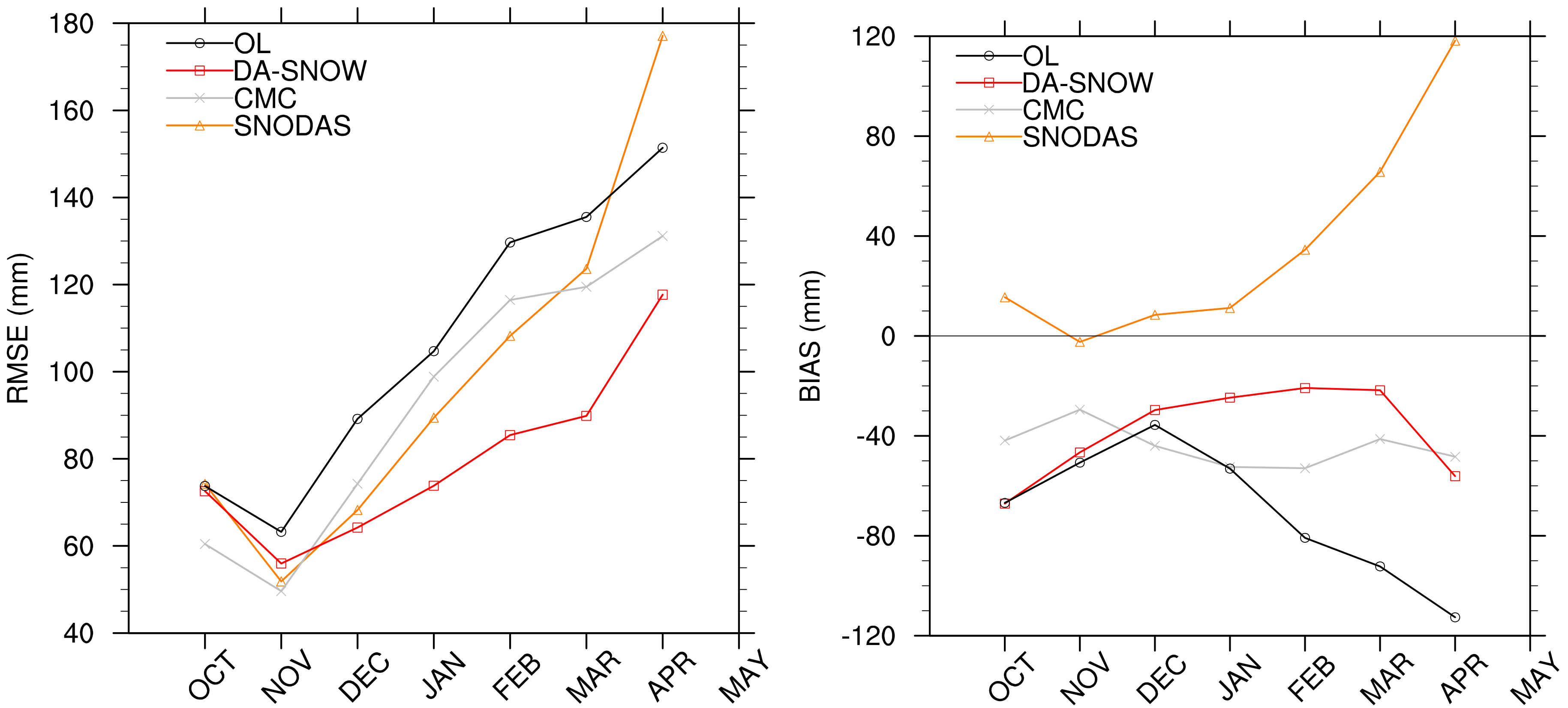




Location of GHCN sites used in the Cressman analysis correction

Snow DA: Evaluation of snow depth fields against GHCN

Average seasonal cycle of snow depth RMSE and bias



	Open loop (no DA)	SNOW-DA	CMC	SNODAS
RMSE (mm)	174.0 +/- 8	**114.0+/- 8**	**158.0+/-8**	**154.0+/- 8**
Bias (mm)	-84.1+/- 8	**-31.6 +/- 8**	**-66.0+/- 8**	**33.9 +/- 8**

Snow DA: Evaluation of the improvements in streamflow simulation

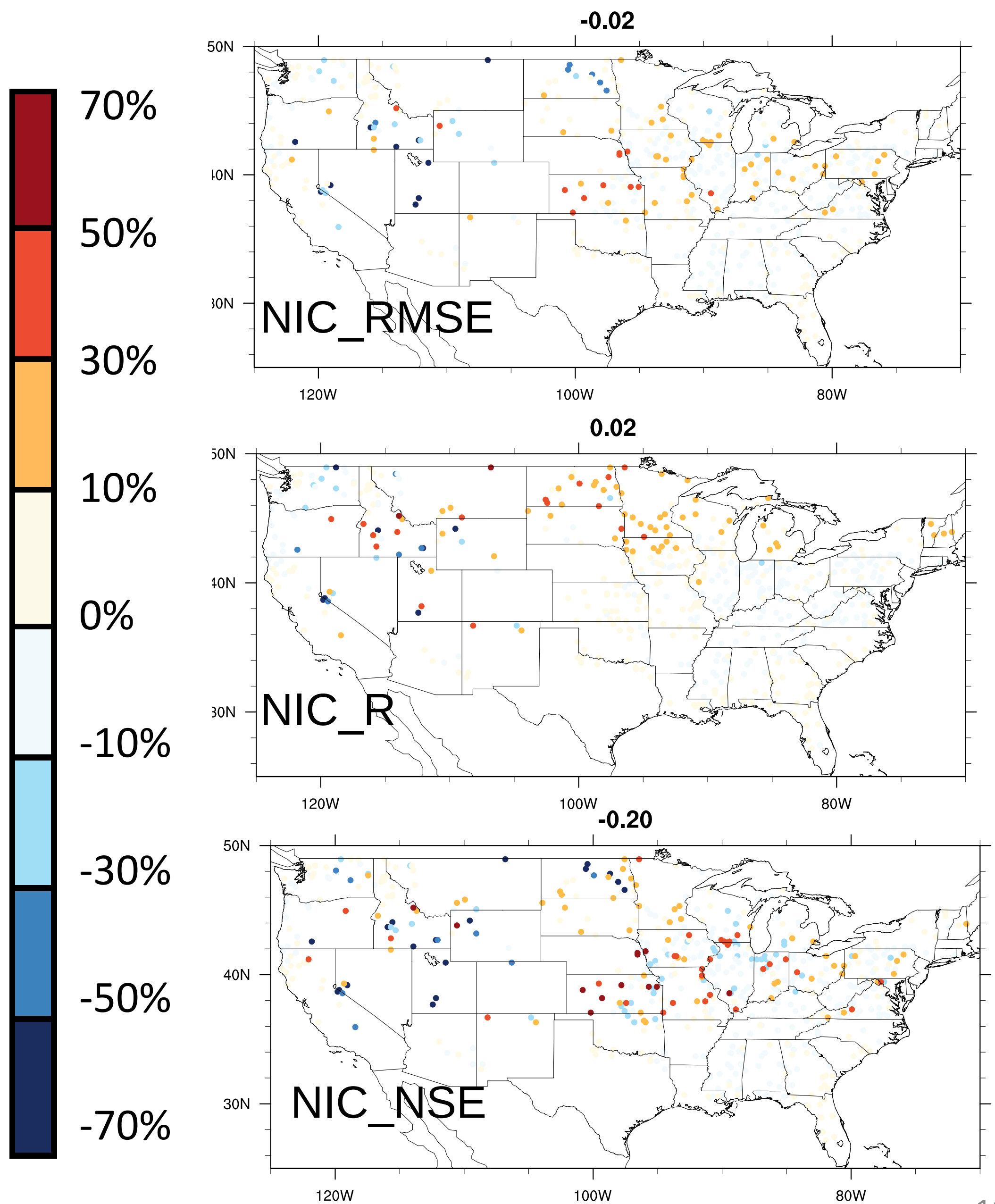
The improvements are expressed using a Normalized Information Contribution (NIC) metric that measures the skill improvement from DA as a fraction of the maximum possible skill improvement

$$NIC_{RMSE} = \frac{(RMSE_o - RMSE_a)}{RMSE_o}$$

$$NIC_R = \frac{(R_a - R_o)}{(1 - R_o)}$$

$$NIC_{NSE} = \frac{(NSE_a - NSE_o)}{(1 - NSE_o)}$$

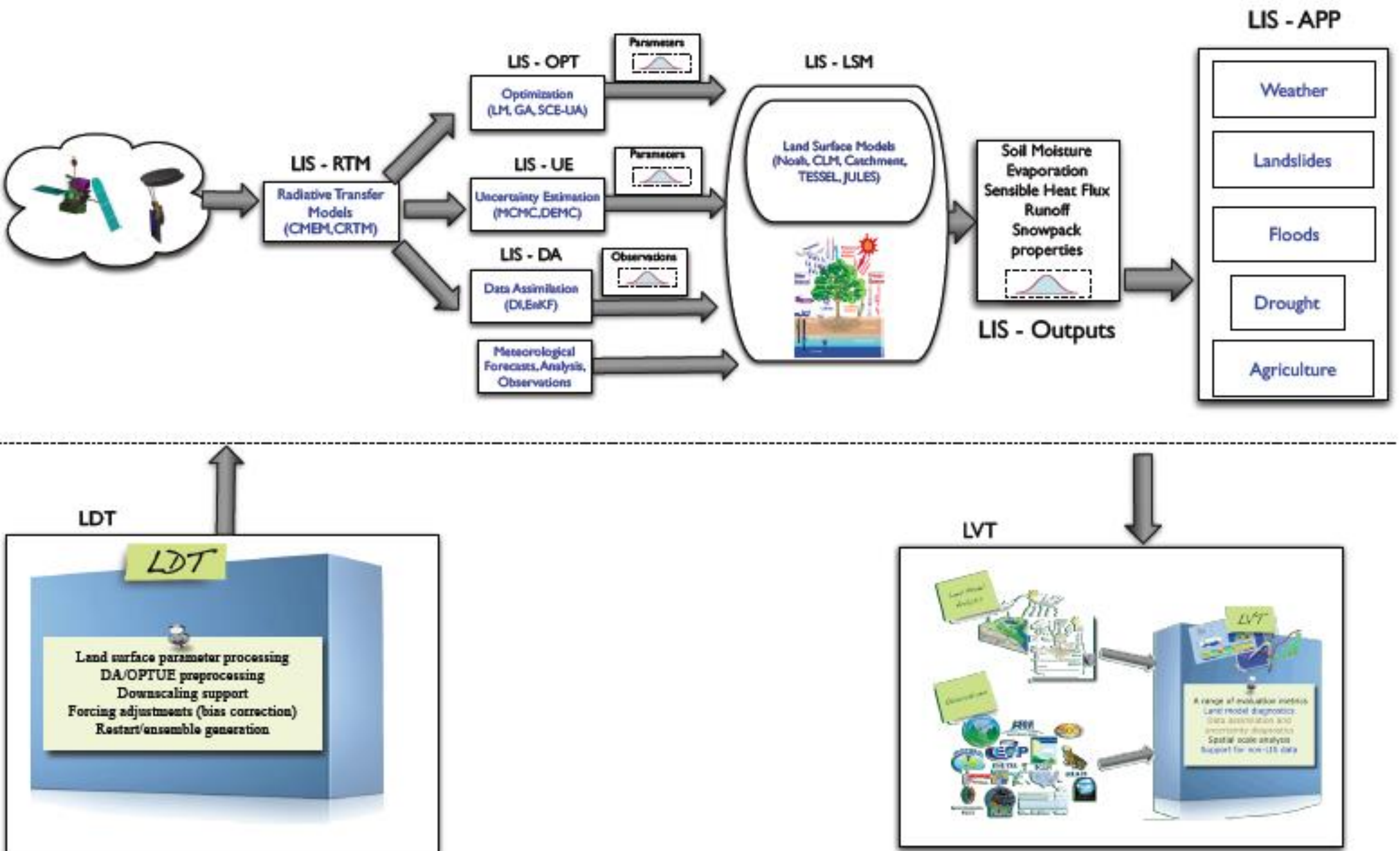
Some improvements in streamflow metrics such as RMSE, R and NSE after snow data assimilation



Additional References

- Kumar, S.V., C.D. Peters-Lidard, J. Santanello, K. Harrison, Y. Liu, and M. Shaw, 2012: Land surface Verification Toolkit (LVT) - a generalized framework for land surface model evaluation, Geosci. Model Dev. , 5, 869--886, doi:10.5194/gmd-5-869-a
- Kumar, S.V., R.H. Reichle, K.W. Harrison, C.D. Peters-Lidard, S.Yatheendradas, J. Santanello, 2012: A comparison of methods for a priori bias correction in soil moisture data assimilation. Wat. Resour. Res., 48 (3), doi:10.1029/2010WR010261.
- Liu, Y., C. Peters-Lidard, S. Kumar, J. Foster, M. Shaw, Y. Tian, and G. Fall, 2013: Assimilating satellite-based snow depth and snow cover products for improving snow predictions n Alaska. Advances in Water Resources, 54, 208-227, doi:10.1016/j.advwatres.2013.02.005.
- Peters-Lidard, C.D, S.V. Kumar, D.M. Mocko, Y. Tian, 2011: Estimating evapotranspiration with land data assimilation systems, Hydrological Processes, 25(26), 3979--3992, DOI: 10.1002/hyp.8387
- Yatheendradas, S., C.D. Peters-Lidard, V.I. Koren, B. Cosgrove, L.G.G. de Goncalves, M.B. Smith, J. Geiger, Z. Cui, J. Borak, S. Kumar, D. Toll, G.A. Riggs and N. Mizukami, 2012 . Distributed assimilation of satellite-based snow extent for improving simulated streamflow in mountainous, dense forests: An example over the DMIP2 western basins. Wat. Resour. Res. DOI:10.1029/2011WR011347

LIS7: New subsystems and toolkits



A formal front-end : Land Data Toolkit (LDT)

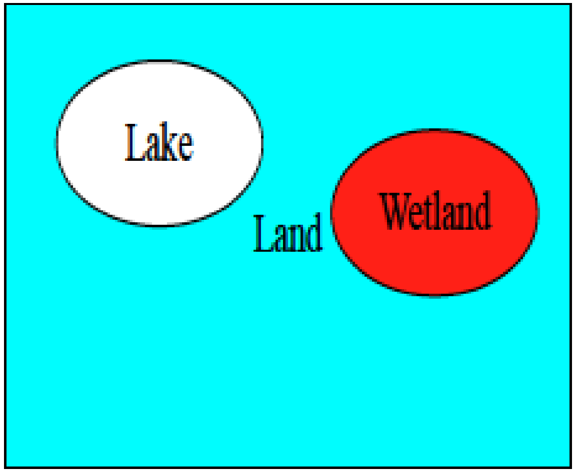
- All static and climatological parameters are handled by LDT and produces a NETCDF input file for LIS.
  - Subsetting/reprojection/aggregation/downscaling of individual (native) parameter data to the LIS target domain
  - Includes cross-checks for ensuring mask consistency, correct selection of parameters, etc.
- Data assimilation preprocessing
  - Generation of climatologies and scaling parameter for bias –correction in LIS-DA
- Restart and ensemble processing
  - Generation of climatological restart files, ensemble restart files

New Land surface models

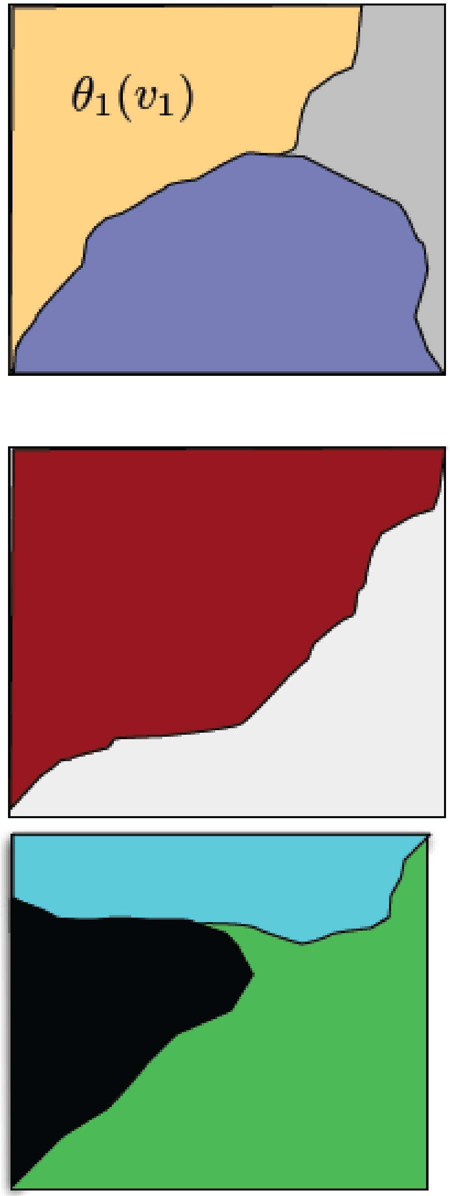
- VIC(4.1.1)
- CABLE
- Catchment (Fortuna 5.2)
- SAC (NLDAS)
- Noah 3.3
- Noah 3.4
- WRSI
- SAC HT-ET
- FASST
- SiB
- JULES

Support for surface models

- LIS7 will allow multiple "surface" models in addition to land surface models.
  - E.g. A domain could consist of land points (running land surface models), lake points (running lake models) and wetland points (running wetland models).
  - LIS would aggregate and "quilt" the outputs from these different model types into a single output structure.
  - LIS7 includes the definition of "patchy domains" that represent the sub-domains that run surface models



Multivariate tiles



LDT tiles the data first:

Categorical variables (landcover, soiltype) :
Compute the percentage of grid points (at 1km) for each type within a coarser resolution gridcell.

Continuous variables (soil fraction, elevation, slope, aspect):
Find the range (max,min) of the variable (at 1km) within a coarser resolution grid cell.
Generate bin sizes based on the prespecified number of bins.
Compute the percentage of grid points (at 1km) that fall into each of the bins. Report the midpoint of the bin as the effective value for that bin.

Forcings

- Forcing structure has been modified to allow for more flexible data transformations
  - Supports overlays
  - Supports forcing ensembles
  - Allows for online bias-correction
- No more “base” and “supplemental” forcings. Everything is a “met” forcing and they can be overlaid in the order in which the user chooses.
  - LIS6 only allows the overlay of supplemental forcings and not baseforcings.

Data Assimilation

- Includes support for smoothing algorithms (e.g. ensemble kalman smoother)
- Support for multiple data products: AMSR-E, SMMR, SSM/I, ECV, ASCAT, GRACE.
- Options for specifying spatially varying error parameters
- Support for radiance assimilation.
- A new implementation of a fast fourier transform and supports horizontal correlations.

Optimization and Uncertainty Estimation

- The new structure allows for the concurrent parameter estimation across different model classes.
  - E.g. parameter estimation of both LSM and RTM parameters against both soil moisture and Brightness temperature observations.

Harrison, K. W., S. V.Kumar, C. D.Peters-Lidard, and J. A.Santanello (2012), Quantifying the change in soil moisture modeling uncertainty from remote sensing observations using Bayesian inference techniques, Water Resour. Res., 48, W11514, doi:10.1029/2012WR012337.

Routing

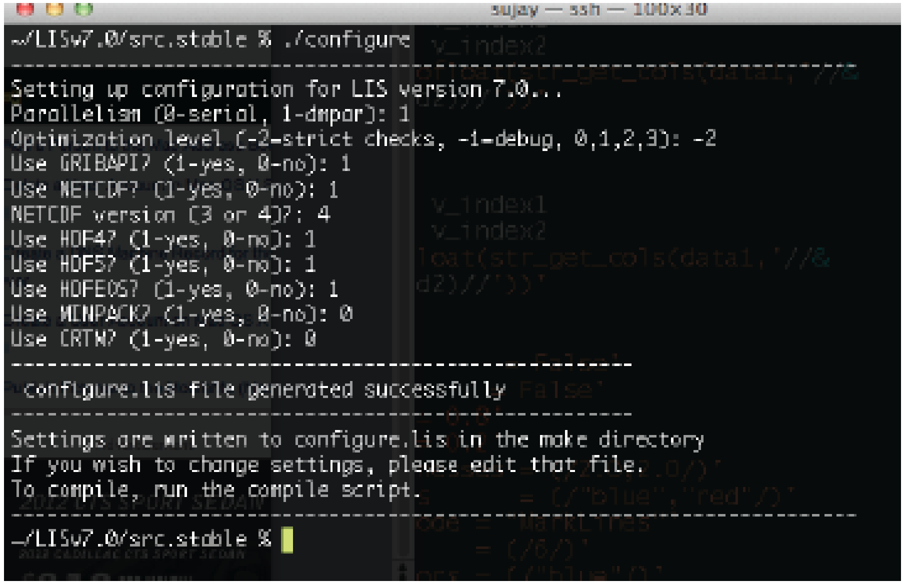
- Will include a suite of routing algorithms
  - Source-to-sink methods: NLDAS router, HYMAP
  - Models that includes lateral transport of soil moisture (and feedback to the model states): NDHMS
- Associated topographical processing will be supported through LDT

NLDAS router: Xia, Y., et al. (2012), Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow, J. Geophys. Res., 117, D03110, doi:10.1029/2011JD016051.

HYMAP router: Getirana, Augusto C. V., Aaron Boone, Dai Yamazaki, Bertrand Decharme, Fabrice Papa, Nelly Mognard, 2012: The Hydrological Modeling and Analysis Platform (HyMAP): Evaluation in the Amazon Basin. J. Hydrometeor, 13, 1641–1665, doi: 10.1175/JHM-D-12-021.1

A new build system

- A perl-based build system
- Prompts the user for the choice of libraries, compile time options (No need to edit the configure.lis or LIS_misc.h files).



Library changes

- Uses ESMF5 series – backward compatibility is ensured.
  - Can use newer releases of ESMF without interface/code changes in LIS.
- Grib – consolidated the use of 3 different grib libraries (NCEP, AFWA, NCAR) with the ECMWF developed grib-api library.
  - Includes a documented F90 API – grib-api
  - Supports both grib1 and grib2
  - No need to distribute libraries with LIS
- LIS7 supports NETCDF4 (and NETCDF3) with options for data compression.
  - NETCDF output follows the CF and COARDS conventions.

Time handling

- LIS7 includes support for 'variable timestepping'
  - Each component (LSM, Forcing, RTM, etc.) sets its own internal timestep.
  - LIS computes the minimum timestep among these components as the timestep for the global clock
  - This enables automatic temporal aggregation of forcings if the LSM is run at a timestep greater than the forcing timestep.
- The use of ESMF-based alarms will be eliminated
  - So that synoptic/monthly/weekly alarm intervals can be handled more easily
  - Resetting/Looping of the global clock can be handled more easily.
  - Allows better nesting support while coupling to WRF.

Configuration

- The C-function tables have been changed from an array structure to a linked list structure
  - This eliminates the need for hardcoded array sizes for C-based function tables
  - This also enables the use of strings as keys to store functions in the C-tables. This leads to a more intuitive lis.config interface:
- Eg:

```
File  Edit  Options  Buffers  Tools  Help
#Overall driver options
Running mode:                                    "retrospective"
Map projection of the LIS domain:                "latlon"
Number of nests:                                  1
Land surface model:                              "NOAH31"
Number of met forcing sources:   1
Blending method for forcings:    "overlay"
Met forcing sources:             "GDAS"
Use elevation correction (met forcing):      "no"
Spatial interpolation method (met forcing):  "bilinear"
Temporal interpolation method (met forcing): "linear"

#Runtime options
Experiment code:                                    '111'  #experiment code
Forcing variables list file:                        ./forcing_variables_v2.txt
```

Misc

- Options for spatial downscaling (slope-aspect correction of radiation, PRISM/WorldClim-based downscaling of precipitation)
- Support for irrigation modeling
- The restart files are written in NETCDF4 formats (as an option)
- Better support for higher compiler optimization levels
- The default 5-level hierarchy of LIS outputs will be changed to a 3-level hierarchy (OUTPUT/MODEL/YEARMONTH)